

CORPORATE CONFIDENTIAL

SPEECHGEAR, INC.

Progress Report

Project

- *Compadre: A Device Independent Voice-to-Voice Language Translator Software Solution*
- SBIR Phase I Topic N01-044
- Contract Number N00014-01-M-0225

Item Number

0001AA: Progress Report

Security Classification

Unclassified

Sponsor

Dr. Joel Davis
Office of Naval Research
Ballston Tower One
800 North Quincy Street
Arlington, VA 22217-5660



Principle Investigator:

Robert Palmquist
SpeechGear, Inc
3454 Circle Bluff Court
Faribault, MN 55021

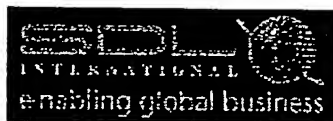


S P E E C H G E A R

Program Partners



VISUAL GOLD



SDL INTERNATIONAL

A r a m e d i A

Software Localization Translation Graphic Design

ARAMEDIA

Compadre: A Device Independent Voice-to-Voice Language Translator Software Solution

*A Phase I SBIR Speech Application Project
for the Office of Naval Research*

*SBIR Call Number N01-044
Contract Number N00014-01-M-0225*



S P E E C H G E A R

**Robert Palmquist
SpeechGear, Inc
3454 Circle Bluff Court
Faribault, MN 55021**

Telephone: 507-664-0974
Email: robert@palmquist.com
Web: www.speechgear.com

A. Project Summary

Technical Abstract:

Mission Statement

To develop and deploy language translation software that is device independent, supports bi-directional translation of multiple languages, produces text transcriptions of spoken conversations and supports translation of text extracted from digital images. This software shall run in both a reduced functionality standalone mode, and by wirelessly connecting to remote servers, a full-function mode. This software shall run on multiple pocketable platforms resulting in a mobile system that is low in cost, easy to use, robust in operation and comfortable to carry and/or wear.

The object of this Phase I research effort is to investigate the scientific, technical and commercial merit and feasibility of the system described in the preceding mission statement. Specifically, the team will investigate design options for the mobile translator system, identify potential applications, and select the best option(s) to pursue in making the design a reality. Four technical areas will be investigated: potential pocketable computing platforms, the operator interface, optical character recognition software and the language translation software. The commercial feasibility of this design will also be investigated. This includes identifying potential applications, languages to be supported, cost, and user requirements such as interface modes and response times. By combining both the commercial and technical elements, a complete definition of successful software and system solutions for pocketable language translation devices will be achieved.

Prototype systems showing device independence will be developed and demonstrated and a final report written documenting the Phase I results and recommendations for follow-on research and development in Phase II. Options are included for incorporating additional language pairs into the system and application specific terminology.

Anticipated Benefits/Potential Commercial Applications of the Research or Development:

Applications include all individuals who require multi-lingual capabilities. The mobile translator will benefit a wide range of individuals including military personnel, airport employees, border patrol and customs agents, police, fire fighters, retail clerks, bank tellers, delivery personnel, phone operators, tourists and any industry that sells, develops or manufactures products to/in global markets or employs individuals that do not speak the native language.

B. Project Status

B.1 Status Overview:

The overall work breakdown structure is provided in Figure 1. For purposes of this report, the project start date is selected at . The actual purchase order has not yet been received in the mail, however a FAX copy of the signed document was provided by Jennifer Schoen on

As is shown in Figure 1, the project is currently ahead of schedule. Prospective Users of the system have been interviewed and the resulting Design Requirements (DR) has been drafted. This document is included in this report as Appendix A. The DR contains the targeted and desired specifications for *Compadre's* overall system performance.

The system is divided into three basic areas: standalone, camera-based and telephone-based operations. These three areas are also listed in order of difficulty, with the standalone mode being the easiest to implement and the telephone based system being most difficult. Progress has been made in each of these three categories. This progress is described in the remainder of Section B.

B.2 StandAlone Mode

In this mode, the PDA/Cellphone (henceforth called a "SmartPhone") will not be required to wirelessly connect to a remote server. The translation capabilities will be primarily bi-directional word look-up. Initially, the interface will be a touchscreen such as is shown in Figure 2. Multiple language pairs will be supported along with a 30,000+ word dictionary.

Six different vendors have been identified for potential teaming partners on developing *Compadre*: AIM, Smart Link, TomTom, Evolutionary Systems, PhatWare and Ectaco. Each of these vendors have provided samples of their current software product, and these are in the process of being evaluated. The templates for this two-part evaluation are provided in Figures 3 and 4. The first template

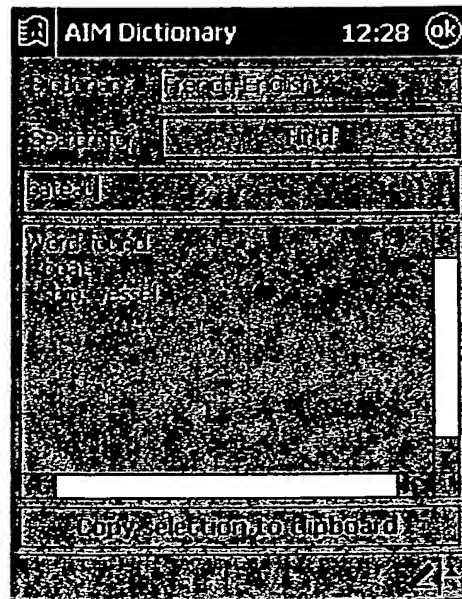


Figure 2: Example of Touchscreen Interface for Stand-Alone Mode

	HPC Translate	Pocket Context	Pocket Language Teacher	Collins Dictionaries	Dictionary	Travel Dictionary
Vendor	PhatWare	SmartLink	Ectaco	TomTom	Evolutionary Systems	AIM
Size of English Dictionary						
Ability to Add User Specific Terminology						
Bi-Directional Capability						
Ease of Use						
Number of Languages Supported						
List of Languages						
List of Supported CPUs						
Additional Comments						

Figure 4: Part 2 of 2 for Evaluating Stand-Alone Translator Products

will be used to test specific word translation capabilities. The second template is used to evaluate overall system capabilities. Based on the results of these tests, two vendors will be selected as partners to continue development activities.

B.3 Camera-Based Mode

The primary means to input text into the SmartPhone for this mode of usage will be a digital camera. A patent application for this capability has been submitted. Such a system is shown in Figure 5. The digital camera will be used to capture an image of the

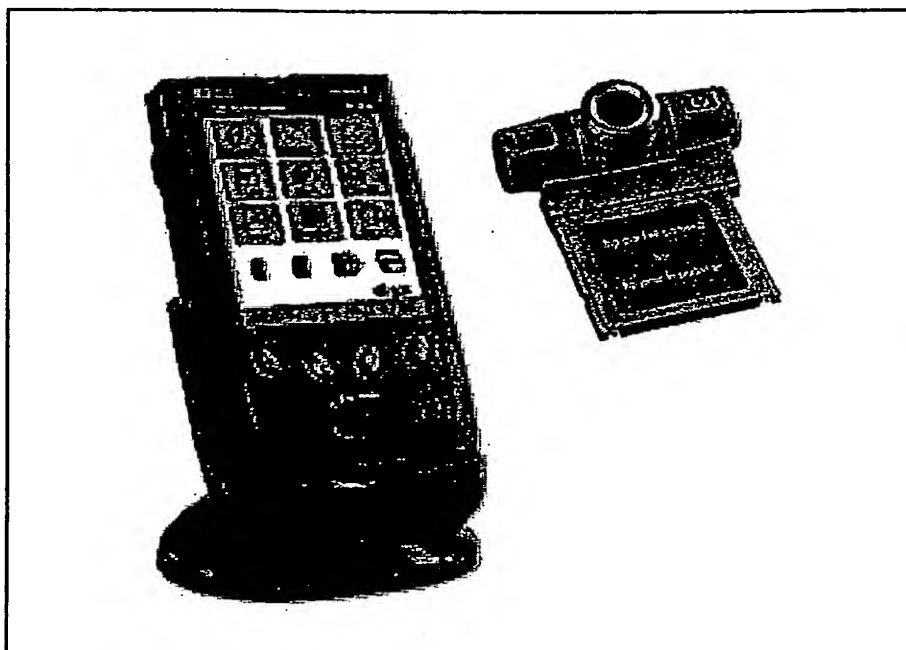


Figure 5: Example of Camera-based System

foreign language. Such a picture is shown in Figure 6¹. Once the desired image is obtained, the SmartPhone will wirelessly connect to a remote server where the image will be processed and the resulting translation sent back to the user. An example of the translated text in the proposed “one-click” GUI is shown in Figure 7. For most applications, this connection will be made using cellular telephones. Because of the limited bandwidth of such a connection, it is important to reduce the overall size of the transmission. Thus, SpeechGear is in the process of evaluating different image compression algorithms. These algorithms will be embedded directly into SpeechGear’s software, and thus will be transparent to the end user. The current plans are to use Visual Gold’s *Imagist* product. This can be viewed at www.visualgold.com. SpeechGear has had initial meetings with Visual Gold and the appropriate NDA’s have been signed. A “Letter of Intent” with respect to the teaming arrangement is in the process of being drafted.

As is shown in Figure 6, a “one-click” GUI is planned. After capturing the image(s), the user will simply select “Translate” and the wireless connection will automatically be established. Note that multiple images can be sent simultaneously using a single click. This is similar to the “Add to Basket” interfaces that are being used at web-based shopping sites. In this approach, items that are selected can be loaded into a virtual basket or cart, and once you are done shopping you can select “Check Out” to purchase all of the items simultaneously. For *Compadre*, multiple images can be selected and entered into the queue, and when the user is ready to connect to the remote server, then

simply selecting the “Translate” button will connect the SmartPhone to the remote server, which in turn will process the images and return the resulting translation. The images will be transmitted back to the user using an HTML format. The users can then scroll

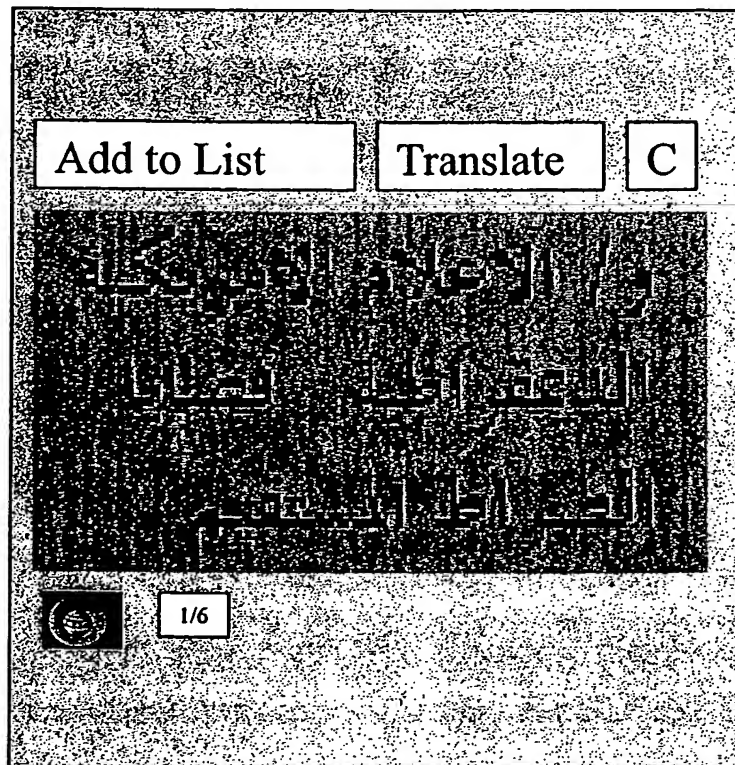


Figure 6: Preliminary Functional Layout of Graphical User Interface – Text Boxes will be Replaced with Icons

¹ Note, since our software is not yet functional, I have no idea what this Arabic text says. If the option is exercised, we will be hiring an Arabic speaking individual to be part of our team.

through these images and save or delete them as is desired. Please note that the actual buttons will be Icons versus text, and thus the look and feel of the resulting GUI will be a substantial improvement over what is shown in the Figures.

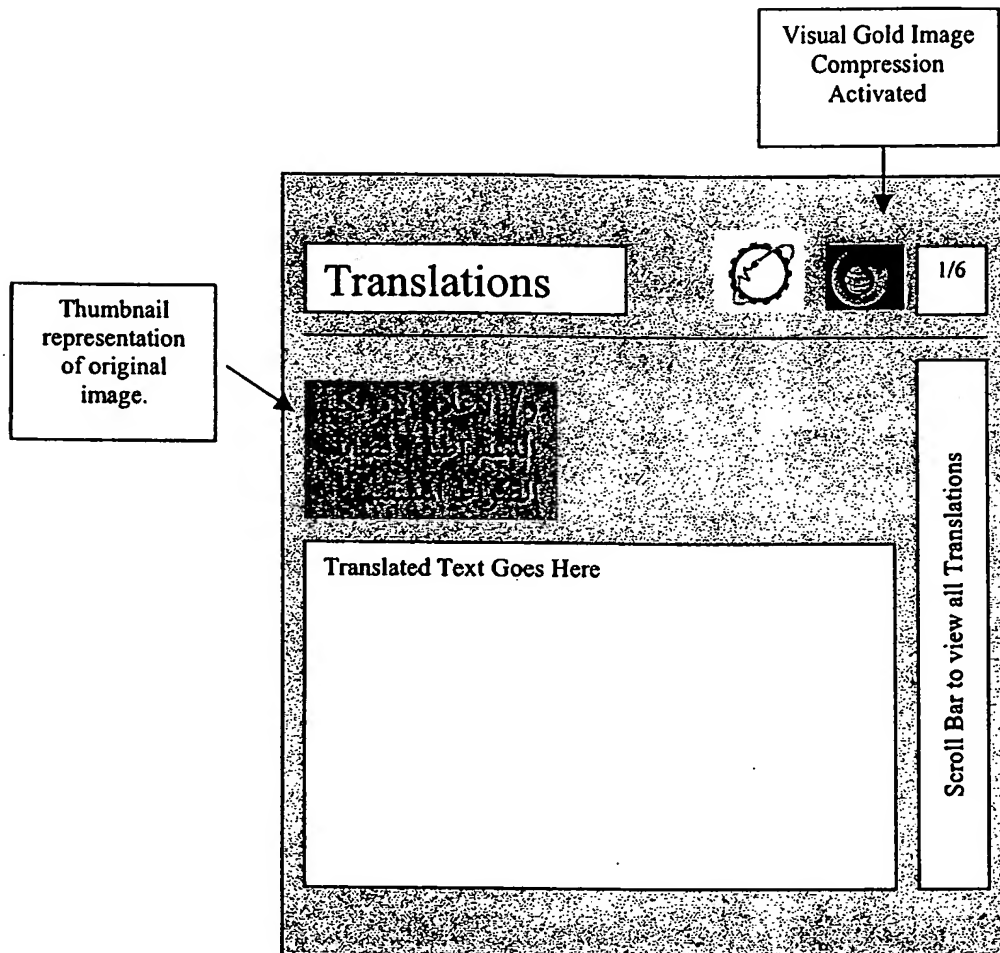


Figure 7: Preliminary Graphical User Interface

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.